The NOAA Operational Model Archive and Distribution System
NOMADS
Fall 2002 Progress Report
November 20, 2002

I. Highlights

- The Climate Diagnostics Center (CDC- Boulder CO) hosted the 2002 NOMADS Team meeting. Over 25 participants attended the meeting (participant list attached). LLNL represented GFDL at the meeting.
- New NOMADS collaborators include: the National Coastal Data Development Center (NCDDC), a sister NESDIS data center. At the request of NODC, NCDDC has agreed to assist in the development of distributed access to ocean related models and data; Unidata has posted Jim Steenburgh, University of Utah as the Unidata university representative to NOMADS; and Chris Hill, MIT and PI of the Earth System Modeling Framework (ESMF) to assist in the advancement of coding standards for GCM models under NOMADS.
- New hires for NOMADS include: NCDC's Ted Smith. Ted will assist with the installation and data configuration of GDS, LAS-CDAT, and the Globus Grid toolkit for the Committee for Earth Observation Satellites (CEOS) Grid collaboration using NOMADS interfaces. NCEP hires NOMADS computer scientist Jun Wang.
- NCEP delivers first Global Spectral Model (GSM) input data set to NCDC. New operational system configuration completed at NCEP that will provide operational data transfer capabilities of NCEP data assimilation files to NCDC.
- NCDC to receive the new NCEP Regional Reanalysis (4/day @32km 1978 to present). This new data set will be provided via NOMADS interfaces. NCEP will dedicate supercomputer resources for this re-run beginning in 1/03.
- Glenn Rutledge presents NOMADS briefings at the NCAR hosted, NSF sponsored Cyberinfrastructure Workshop; and at the ECMWF 10th Workshop on High Performance Computing in Meteorology in Reading UK, and CEOS-Gird meetings.
- European interest remains high NOMADS collaborator Bryan Lawrence, Head, British Atmospheric Data Center (BADC) has recently installed NOMADS compatible Web services.

II. Upcoming Events

- NCDC is finalizing the NOMADS access interfaces to the NCEP RUC model, (awaiting Systems Branch new Firewall), and developing an informational Web page containing NOMADS overview, available data sets, status reports, participants, templates to join, by 1/03.
- The NOMADS team has agreed to develop a Steering Committee to develop short and long range goals for NOMADS and develop a "Federation" of partners to provide for Agency attribution for participants.
- A NOMADS requirements process is underway, along with the development of the NOMADS Program Plan to be completed by 1/03.

II. NOMADS Yearly Meeting Summary

All NOMADS collaborators were in agreement that NOMADS has seen exceptional national and international recognition and growth. This momentum is seen by invitations to use NOMADS services in projects such as CEOS, CEOP, e-science NERC-DataGrid, and the NSF Cyberinfrastructure efforts. NOMADS services to these organizations is in the form of a package of freely available software and a framework which when added to an existing agency/project's computer/web services allows the distribution of data sets in a near format independent method.

A NOMADS participant, in this case the agency or project disseminating the data set, controls the integrity and documentation of datasets. The display and scientific manipulation of data, is accomplished by the client (data receiver), who can elect to use their own public or commercial display package. A NOMADS server has the ability to dissect data sets and recompose fields and make calculations as desired by the client. Web based ftp can also "slice and dice" data sets according to users needs as part of the NOMADS framework.

The NOMADS collaborators agree that the continuing success of NOMADS depends upon some level of NOAA base funding to support 1) a steering committee; 2) a NOMADS software control and NOMADS Web team; and 3) funding for OPeNDAP (see below). NOMADS efforts to attract funding need to include funding directed to OPeNDAP to help to address NOMADS priorities. It was noted however that NOMADS as an agreement between agency participants act as a leveraging process to obtain local funding and insure a common data delivery framework. In fact CDC, GFDL, NCAR, NCDC, NCEP, LLNL and PMEL have already directed local resources in support of NOMADS.

III. A Federation of Distributed Data Services

The team agrees that the development of a Unified or "Federated" Web page needs to be developed. This unified page will provide for attribution for Agency participation. After appropriate inter-agency coordination and agreement, the Federation would be created such as the National Environmental Distributed Data Services- "A Collaborative Federation for Environmental Services".

A unified NOMADS Web Page that can act as a guide for participants to present their services needs to be developed over the next FY. Each participating Center is currently working to develop its own Web interfaces, which is site specific, and developed to best serve their own unique data center requirements. However, these interfaces are difficult to use and understand, and the new user will not interact with more advanced NOMADS clients such as CDAT, LAS or GrADS, Ferret, or IDL. Currently, GFDL's NOMADS usage statistics indicate that the most extensive use of NOMADS interfaces continue to be common ftp data transfers. A common NOMADS Web interface, with simple data retrieval methods and a User's Guide along with information concerning the NOMADS program, participants, and host templates for becoming a user- is now required.

IV. GRID Services

Another goal discussed for FY03 and beyond will be the implementation of open source Grid-based technologies into NOMADS. Grid services such as the Hierarchical Storage Resource Manager (HRM) may be implemented to interface with HPSS and other mainframe archive systems for retrospective access to NOMADS data. Other near real-time Grid services such as "Grid-FTP" will allow for high volume point-to-point data transfers between Centers and Laboratories.

NOMADS has been selected as one of 4 prototype applications to reside on the new CEOS-Grid. Globus toolkits, security certificates and eventually HRM resources will be installed and prototyped on NCDC servers that will be linked to NASA, DOE, USGS, and the European Space Agency (ESA). The NESDIS Chief Information Officer (CIO) is supporting this effort at NCDC.

V. Systems Development

The team has reached agreement that each participating NOMADS site would serve their data both through OPeNDAP (incl. GDS) and through LAS. Each site may have a unique LAS back end that incorporates their in-house tools (CDAT, NCL, IDL, MatLab, Ferret, GrADS, etc.). The team will build a unified Web entry point to all of the data using LAS "sister server" collaboration, where OPeNDAP makes comparisons at the binary level possible (work remains on this piece). A CDAT back-end in the unified server would provide a mature model inter-comparison metrics framework. A range of OPeNDAP-enabled desktop tools would be available for detailed, custom analysis: Matlab, IDL, GrADS, Ferret, Excel. Our goal is to migrate the entire framework to work within the GRID (Earth System Grid, CEOS-Grid).

The data format neutral "glue" that holds the effort together is OPeNDAP (formally known as DODS). OPeNDAP is a software framework, and consortium that is a viable project with project resources. NOMADS uses this framework and use of these data conventions (_cf, NetCDF, or any OPeNDAP enabled client) greatly advances (but does not ensure) for cross discipline interoperability (including ocean and in-situ) between the climate and ocean communities. The team agreed that the current level of interoperability was sufficient for the short-term, but that steps need to be taken to ensure for use of new data forms in the future (e.g., satellite and radar). To work toward the best use of limited ESDIM funding, Steve Hankin (PMEL) has submitted a 2003 ESDIM proposal that establishes an OPeNDAP Steering committee, comprised of representatives from each NOAA line office that would provide resources to OPeNDAP efforts that advance the NOMADS distributed data access philosophy. This same OPeNDAP Steering Committee could be expanded for NOMADS to include other NOMADS participating government agencies (e.g., DOE, NASA), institutions (e.g., NCAR, COLA), and international partners (e.g., BADC, eScience). The formation of a NOMADS Steering Committee and how it will operate will be advanced during FY03.

In the long term an "open source" modular design is being considered for integration into the NOMADS core infrastructure. This will initially include the CDAT tools that are Python based. In this way, NOMADS users can have access to discipline independent

libraries of user contributed modules of equations and formulas, statistical routines, and other tools as developed in other scientific disciplines such as Physics, Mathematics, or Medicine.

VI. International Data Management Meeting

Regarding the need for an international data management meeting, the team recommends that a meeting at the technical level would be valuable; the goals are currently too ill defined at this time and this needs further work before the meeting is requested. However it was also noted that the current NOMADS OPeNDAP framework allows for significant interoperability between the ocean and atmospheric communities for the short term. It was suggested that either NOAA or several NOAA Line Offices initiate planning sessions now for a potential Fall 2003 meeting.

VII. Next Steps

1. The next steps are to develop the NOMADS Program Plan with a Users Guide and to develop a Federation of participants. This Federation will host the unified NOMADS web site that has information about NOMADS, its members, how to become a member, listing of quarterly reports including technical specifications and Agency attribution.
2. Implement a NOMADS Steering Committee.
3. NCDC must include it's Web interface staff with NOMADS collaborators to ensure NCDC operations is actively involved in this rapidly growing project, and collaboration in terms of technical data distribution efforts NOAA-wide and especially within NCDC.
4. NCDC must develop science-based partnerships with GFDL, NCAR and elsewhere, outside the NOMADS framework in order to then develop NOMADS partnerships using NOMADS distributed data. NOMADS is data distribution/access; whereas science based activities must begin either separately or in concurrently with the NOMADS access or data distribution philosophy. In order to achieve these new collaborations, NCDC will advance the use of these technologies within its Scientific Services Division.
5. NOMADS will work toward the implementation of Grid technologies within the Earth System Grid (ESG), and the CEOS DataGrid. This will advance NCDC's understanding of grid technologies including a functionality called the Hierarchical Resource Management (HRM) that would interface NOMADS spinning disks with HPSS-based mass store archives. HPSS is currently in use within the NCDC main archive.

End of Report
G. Rutledge
11/20/2002

NOMADS Collaboration Work Strategy
October 24, 2002

(With input provided from NCDC, NCEP, GFDL, CDC, PMEL, NCAR and LLNL, and the NOMADS participants at the yearly workshop in Boulder)

1. **What are the Goals of NOMADS?**
    a. **Short Term Goals**
        i. Development of a NOMADS Program Plan and Operating Procedures and Project Milestones for FY03. Action: Rutledge. Due (for team review): 1/03
        ii. Work to secure resource funding. Action: NCDC, NCEP, GFDL, CDC. Due: on-going
        iii. Determine relationship between NOMADS and DOE's Earth System Grid (ESG). Action: Rutledge, Williams.
        iv. Finalize NCDC's GDS system and get NOAAPort data on line. Action: Rutledge. Due 3/03
        v. Complete BUFR station data handling capability for GrADS. Action: Doty (COLA) Due: 12/02
        vi. Develop an integrated Web presence for NOMADS: a unified front-end Web Page for NOMADS. Action: Hankin, Williams, Kerr, Rutledge. Due: 3/03
        vii. Work with existing efforts (GODAE, GCOS, BADC, et al.,), to ensure for cross disciple inter-operability and plan for international data management meeting (see Section 4)

    b. **Long-Term Goals**
        i. Further develop NOMADS and CEOS-Grid and ESG partnerships
        ii. Secure funding resources for all core NOAA participants: NCEP, GFDL, NCDC, PMEL, and CDC.
        iii. Integrate Python-based technologies (CDAT, Globus) into NOMADS in parallel with standard back-end services (GrADS, LAS, OPeNDAP).
        iv. Develop Steering Committee or integrate NOMADS philosophy into existing NOAA operational structures (CIO's?).

2. **NOMADS Program Plan Outline Discussion**
    Note: These items were discussions held during and after the NOMADS Workshop with key NOMADS participants.

    a. **What is NOMADS?**
        i. NOMADS is an agreement between agencies, who participate, to have common data and observation distribution software, using a format independent and description methodology, and for a unified documentation and organizational framework for data distribution.

ii. NOMADS provides a forum to plan and organize these structures for use by university, federal agencies and organizations as participants (servers of data) or as clients (users of data).

iii. NOMADS is a platform for these agencies to obtain support for the dissemination of their data sets and to make use of these data sets. NOMADS participants will provide their own resources to accomplish this. A resource means the hardware and the programming staff to execute their plans. Clients should invest in the training it takes to use a web page client, like ftp2u, or client software like GrADS, IDL, MatLab, Ferret, Live Access Server, idv, etc.

iv. Eventually, NOMADS will integrate Grid based technologies- the Open Source Python language, to allow easy access to large data sets (time series of multiple files (e.g., 5 years of u/v from the Eta model) in a modular programming language. Through choices of user services- GDS, LAS, OPeNDAP, HTTP, ESG, and CEOS-Grid, these choices will allow for model, observational and other (i.e., satellite data in HDF format, and OpenGIS data structures), data using LLNL developed Climate Data Analysis Tools. Under Python, and the modular Open Source design, other discipline specific routines would become available to the NOMADS user (e.g., high energy plasma physics statistics routines would be "plug and play" for the NOMADS user.

v. NOMADS is a user of the OPeNDAP framework. OPeNDAP is a stand-a-lone server. The DAP protocol (OPeNDAP's core) has no dependencies to the HTTP library or any other communications library. ESG has been working on a new Point-to-Point Transfer (PPT) communication library for OPeNDAP. This lightweight library allows easy communications with OPeNDAP servers so data can be transferred from OPeNDAP servers to GridFTP servers. GridFTP uses Globus Grid technology to transfer large amounts of data securely and quickly.

b. What is the Real-Time NWP NOMADS Server?

i. The NCEP Real-time NOMADS (RT-NOMADS) prototype project serves real time operational data only. NCDC is the operational archive focal point and project leader for NOMADS and holds data sets older than real time. The NCEP RT-NOMADS server sends NCDC data for archival. The NCDC NOMADS goal is to save model (run history) data but budget realities provide that the archives can save initial conditions and observations sufficient to restart NCEP models to reconstruct model run history. A NOMADS goal at NCDC is to add to this archive, the NCEP operational run history forecasts as soon as sufficient storage is funded. Other NOMADS participants serve their own data sets. The NOMADS participant data sets will be connected by various search engines now under development at UCAR (THREDDS), NCAR, and NASA. The RT-NOMADS project

is a prototype with the goal to produce an Operational Specification which if built is planned to be located in the operational data distribution component of NWS.

c. What will the NOMADS Program Plan do?

i. The NOMADS Program Plan will outline roles and responsibilities across the participants. It will contain a User's Guide and a template for others to establish and serve their own data sets under the NOMADS umbrella. The Plan will contain the list of data conventions currently in use and how to deal with intercomparison of different data types and appendices of currently available data.

d. What are the data conventions in NOMADS? Is there any convergence as to a "standard"?

i. COARDS
ii. _cf
iii. GRIB
iv. BUFR
v. GRIB2
vi. FGDC
vii. HDF
viii. If it does OPeNDAP it's NOMADS
ix. Non-COARDS compliant data sets can be accessed
x. ascii (using ftp) services available.
xi. _cf appears to be gaining popularity across the climate community and NOMADS will embrace this form as (one of its) "standard".

e. What data forms do we expect to use in the future?

i. OpenGIS
ii. others tbd

f. Who are the core NOMADS participants and data suppliers?

i. NCDC
ii. PMEL
iii. NCEP
iv. GFDL
v. CDC
vi. NCAR
vii. PCMDI
viii. USGODAE
ix. COLA
x. NASA/GCMD (metadata and data locations)
xi. International (long term)

g. Who are the potential data suppliers?

i. Other data centers (NESDIS, NWS, NOS, NMFS, Unidata, etc.)

ii. CLASS
iii. IOOS
iv. GOOS
v. GCOS
vi. CEOP
vii. CEOS
viii. Grids: NERC- Data Grid, e-science, IPG, ESG
ix. Other International:
    1) GODIVA (Keith Haines) Reading
    2) ESMF (modeling code framework)
    3) PRISM (NOMADS like activity in Europe)
    4) ESG (Dean Williams)
    5) BOM (Australia thru GODAE)
    6) BADC (Bryan Lawrence)
    7) CEOS-Grid (Rutledge, Diamond)
    8) Project CEOP (Japan)
    9) Climateprediction.net

h. What are the core Application Programming Interfaces (API's)?

i. JAVA-OPeNDAP
ii. C++-OPeNDAP
iii. NetCDF
iv. THREDDS
v. Python and Climate Data Analysis Tools (CDAT)
vi. Eventually include the work of the BADC for Grid based direct GRIB and BUFR interface.

i. What are the core NOMADS communications protocols?

i. OPeNDAP
ii. HTTP
iii. FTP
iv. LDAP
v. THREDDS
vi. XML
vii. Globus/CEOS-Grid/ESG

3. What is the relationship between NOMADS and the "grid" especially with DOE's Earth System Grid?

a. The long-term viability of NOMADS will be directed toward the collaborative nature of the current NOMADS partners, and future expected projects particularly the Earth System Grid, Ceos-Grid, and the British Atmospheric Data Center (BADC). BADC will provide access to the European community suite of data including ECMWF, Hadley and eventually UKMET. Australia and Japan are already participating under USGODAE, and CEOP (not yet a firm CEOP commitment). The vision currently taking shape is that NOMADS will be the umbrella under which ESG, CEOS-Grid, NERC

DataGrid (UK), and other efforts will operate. Under an open source Python-based operating environment, NOMADS users will be able to use whatever tool they so desire and the NOMADS-ESG will allow users to intercompare data of differing formats and from differing locations using either thin-client (Web browser), or up to thick-clients (host services using CDAT, and/or Internet2). Once we have this framework, day-to-day maintenance will be secured 1st via NOAA level initiatives then on to NOAA core mission base funding.

4. **Is there a need for an international data management meeting?**
    a. All attendees are in agreement that an international data management meeting needs to be conducted at the technical level. The team recommends that the meeting be planned/organized at the NOAA or line-office level. Summer or Fall of 2003 at the earliest was suggested based upon other commitments and programs. Action (to formulate planning sessions): Rutledge/Hankin/Karl/N. Smith

5. **How do we best leverage existing and on-going work for a unified NOMADS presence?**
    a. CDC, GFDL, and others, specifically the IRI at Columbia, and PCMDI at LLNL have extensive experience with model and in-situ Web services and the NOMADS group need to better leverage from their experiences. Currently, the back-end services (GrADS, Live Access Server, DODS) are well defined, with software development groups leading the way. NOMADS will not substantially modify these groups' efforts. What is needed, and where NOMADS can be most effective during the next year is the development of easy to understand, front-end GUI / Web interfaces to NOMADS services. Currently each participation lab or Center is developing it's own Web interfaces. The interface should be modular so that a participating Center can pick and choose which NOMADS services they desire for internal use or public access.
    b. Another necessary functionality required for NOMADS and not currently available is the NOMADS spinning disk to mainframe archives (HDSS, Mass Store at NCAR, SP's at NCEP, etc.). This activity will be initiated at NCDC during FY03.
    c. The development of a steering committee to direct the overall objectives, from an international perspective for data interoperability is recommended. Who leads such a group was not determined. The NOMADS team reiterated the need for such a committee to have some form of resource allocation capability in order to prioritize and direct resources as required.

The NOAA Operational Model Archive and Distribution System (NOMADS)
Fall Workshop Attendees
(Listed alphabetically by organization)

BADC
Bryan Lawrence
Head, British Atmospheric Data Centre
Rutherford Appleton Laboratory
Chilton, Didcot, OX11 0QX, U.K.
www.badc.rl.ac.uk
B.N.Lawrence@rl.ac.uk

CDC
Donald Mock
Executive Director
NOAA-CIRES Climate Diagnostics Center
R/CDC
325 Broadway
Boulder, CO80305
drm@cdc.noaa.gov

Michael Alexander
NOAA-CIRES Climate Diagnostics Center
R/CDC1
325 Broadway
Boulder, CO 80305
maa@cdc.noaa.gov

Don Hooper (Workshop Host)
NOAA-CIRES Climate Diagnostics Center
R/CDC1
325 Broadway
Boulder, CO 80305
hoop@cdc.noaa.gov

Roland H. Schweitzer
NOAA-CIRES Climate Diagnostics Center
R/CDC1
325 Broadway
Boulder, CO 80305
rhs@cdc.noaa.gov

CDC (cont.)
Jeffrey Whitaker
NOAA-CIRES Climate Diagnostics Center
R/CDC1
325 Broadway
Boulder, CO 80305
jsw@cdc.noaa.gov

CGAM
Eric Guilyardi
Centre for Global Atmospheric Modeling
Dept. of Meteorology PRISM Project
University of Reading
Reading RG6 6BB UK
ericg@met.reading.ac.uk
(attending for):
Reinhard G. Budich Director, PRISM
Max Planck Institute for Meteorology
Bundesstr.55
D-20146 Hamburg Germany
budich@dkrz.de

COLA
Brian Doty
The Center for Ocean-Land-Atmosphere Studies
Institute for Global Environment and Society
4041 Powder Mill Road, Suite 302
Calverton, MD 20705-3106
doty@cola.iges.org
Joe Wielgosz
Center for Ocean-Land-Atmosphere Studies Institute for Global Environment and Society
4041 Powder Mill Road, Suite 302
Calverton, MD 20705-3106
joew@cola.iges.org

FNMOC
Phil Sharfstein
Fleet Numerical Meteorology and Oceanography Center
US Navy GODAE Program Manager
Monteray, CA
phil.sharfstein@metnet.navy.mil

GFDL
(Dean Williams briefing for)
Ron Stouffer
Geophysical Fluid Dynamics Laboratory
NOAA/OAR/GFDL
Geophysical Fluid Dynamics Laboratory
Princeton, NJ 08542
Ron.Stouffer@noaa.gov

GMU
Ruixin Yang
Center for Earth Observing and Space Research (CEOSR)
Department of Physics
George Mason University
Fairfax, VA 22030-4444
yang@yang.gmu.edu

LLNL/PCMDI
Mike Fiorino
Program for Climate Model Diagnosis and Intercomparison
DOE/LLNL/PCMDIL?264
Lawrence Livermore National Laboratory
Livermore, CA 94550
fiorino@llnl.gov

Dean Williams
Program for Climate Model Diagnosis and Intercomparison
DOE/LLNL/PCMDIL?264
Lawrence Livermore National Laboratory
Livermore, CA 94550
williams@pcmdi.llnl.gov

NCAR
Don Middleton
National Center for Atmospheric Research
1850 Table Mesa Drive
Boulder, CO 80305
don@ucar.edu

Dave Brown
National Center for Atmospheric Research
1850 Table Mesa Drive
Boulder, CO 80305
dbrown@scd.ucar.edu

NCDC
Glenn K. Rutledge
NOMADS Principle Investigator
NOAA National Climatic Data Center
151 Patton Ave
Asheville, NC 28801-5001
Phone:(828) 271-4097
Glenn.Rutledge@noaa.gov

NCEP
Jordan Alpert
NOAA National Weather Service
National Centers for Environmental Prediction
5200 Auth Rd.
Camp Springs, MD 20746
jalpert@ncep.noaa.gov

NGDC
Mark Ohrenschall and John Cartwright for Ted Habermann
NOAA National Geophysical Data Center
E/GC 325 Broadway
Boulder, Colorado USA 80305-3328
Ted.Habermann@noaa.gov

PMEL
Steve Hankin
Pacific Marine Environmental Laboratory
7600 Sand Point Way NE
Seattle, WA 98115-0070
hankin@pmel.noaa.gov

Don Denbo
Pacific Marine Environmental Laboratory
7600 Sand Point Way NE
Seattle, WA 98115-0070
Donald.W.Denbo@noaa.gov

UCAR Unidata
Ben Domenico (Day 1)
Linda Miller
John Caron
Ethan Davis (NOMADS POC)
Unidata University Corporation for Atmospheric Research
P.O. Box 3000
Boulder, CO 80307-3000
Ethan.Davis@unidata.ucar.edu

URI
Peter Cornillon
Graduate School of Oceanography
University of Rhode Island
Narragansett RI 02882
pcornillon@gso.uri.edu

James Gallagher
Graduate School of Oceanography
University of Rhode Island
Narragansett RI 02882
jgallagher@gso.uri.edu